Genetic genealogical models in rare event analysis

Frédéric Cérou, Pierre del Moral, François Le Gland, Pascal Lezaud

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Genetic Genealogical Models in Rare Event Analysis

Frédéric Cérou , Pierre Del Moral , François Le Gland , Pascal Lezaud

N°5878

Avril 2006

Systèmes numériques

ISSN 0249-6399 ISRN INRIA/RR--5878--FR+ENG

Genetic Genealogical Models in Rare Event Analysis

Frédéric Cérou* , Pierre Del Moral† , François Le Gland‡ , Pascal Lezaud§

Systèmes numériques
Projet Aspi

Rapport de recherche n ° 5878 — Avril 2006 — 29 pages

Abstract: We present in this article a genetic type interacting particle systems algorithm and a genealogical model for estimating a class of rare events arising in physics and network analysis. We represent the distribution of a Markov process hitting a rare target in terms of a Feynman–Kac model in path space. We show how these branching particle models described in previous works can be used to estimate the probability of the corresponding rare events as well as the distribution of the process in this regime.

Key-words: interacting particle systems, rare events, Feynman-Kac models, genetic algorithms, genealogical trees

(Résumé : tsvp)

This work was partially supported by CNRS, under the projects *Méthodes Particulaires en Filtrage Non–Linéaire* (project number 97–N23 / 0019, *Modélisation et Simulation Numérique* programme), *Chaînes de Markov Cachées et Filtrage Particulaire* (*MathSTIC* programme), and *Méthodes Particulaires* (AS67, *DSTIC Action Spécifique* programme), and by the European Commission under the project *Distributed Control and Stochastic Analysis of Hybrid Systems (HYBRIDGE)* (project number IST–2001–32460, *Information Science Technology* programme).

* IRISA/INRIA, Campus de Beaulieu, 35042 RENNES Cédex, France. `Frederic.Cerou@inria.fr`
† Laboratoire J.A. Dieudonné, Univ. Nice, Sophia-Antipolis, Parc Valrose, 06108 Nice, Cédex 2, France. `delmoral@math.unice.fr`
‡ IRISA/INRIA, Campus de Beaulieu, 35042 RENNES Cédex, France. `legland@irisa.fr`
§ Centre d'Etudes de la Navigation Aérienne, 7 avenue Edouard Belin, BP 4005, 31055 TOULOUSE Cédex 4, France. `lezaud@cena.fr`

Unité de recherche INRIA Rennes
IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex (France)
Téléphone : 02 99 84 71 00 - International : +33 2 99 84 71 00
Télécopie : 02 99 84 71 71 - International : +33 2 99 84 71 71

Modèles Génétiques Généalogiques pour l'Analyse d'Evénements Rares

Résumé : Nous présentons dans cet article un algorithme de particules en interaction de type génétique et un modèle généalogique pour estimer une classe d'événements rares provenant de la physique et de l'analyse des réseaux. Nous exprimons la loi d'un processus de Markov qui atteint un événement rare comme un modèle de Feynman–Kac dans l'espace des trajectoires. Nous montrons comment ces modèles de branchement de particules décrits dans des travaux précédents peuvent être utilisés pour estimer la probabilité d'événements rares, ainsi que la loi du processus dans ce régime.

Mots-clé : systèmes de particules en interaction, événements rares, modèles de Feynman-Kac, algorithmes génétiques, arbres généalogiques

1 Introduction

Let $X = \{X_t\,,\, t \geq 0\}$ be a continuous–time strong Markov process taking values in some Polish state space S. For a given target Borel set $B \subset S$ we define the hitting time

$$T_B = \inf\{t \geq 0 \,:\, X_t \in B\}\;,$$

as the first time when the process X hits B. Let us assume that X has almost surely right continuous, left limited trajectories (RCLL), and that B is closed. Then we have that $X_{T_B} \in B$. In many applications, the set B is the (super) level set of a scalar measurable function ϕ defined on S, i.e.

$$B = \{x \in S \,:\, \phi(x) \geq \lambda_B\}\;.$$

In this case, we will assume that ϕ is upper semi–continuous, which ensures that B is closed. For any real interval I we will denote by $\mathbb{D}(I,S)$ the set of RCLL trajectories in S indexed by I. We always take the convention $\inf\emptyset = \infty$ so that $T_B = \infty$ if X never succeeds to reach the desired target B. It may happen that most of the realizations of X never reach the set B. The corresponding rare event probabilities are extremely difficult to analyze. In particular one would like to estimate the quantities

$$\mathbb{P}(T_B \leq T) \quad \text{and} \quad \mathrm{Law}(X_t\,,\, 0 \leq t \leq T_B \mid T_B \leq T)\;, \tag{1.1}$$

where T is either

- a deterministic finite time,
- a $\mathbb{P}$–almost surely finite stopping time, for instance the hitting time of a recurrent Borel set $R \subset S$, i.e. $T = T_R$ with

$$T_R = \inf\{t \geq 0 \,:\, X_t \in R\} \quad \text{and} \quad \mathbb{P}(T_R < \infty) = 1\;.$$

The second case covers the two “dual” situations.

- Suppose the state space $S = A \cup R$ is decomposed into two separate regions A and R. The process X starts in A and we want to estimate the probability of the entrance time into a target $B \subset A$ before exiting A. In this context the conditional distribution (1.1) represents the law of the process in this "ballistic" regime.
- Suppose the state space $S = B \cup C$ is decomposed into two separate regions B and C. The process X evolves in the region C which contains a collection of "hard obstacles" represented by a subset $R \subset C$. The particle is killed as soon as it enters the "hard obstacle" set R. In this context the two quantities (1.1) represent respectively the probability of exiting the pocket of obstacles C without being killed and the distribution of the process which succeeds to escape from this region.

In all the sequel, $\mathbb{P}(T_B \leq T)$ will be of course unknown, but nevertheless assumed to be strictly positive.

The estimation of these quantites arises in many research areas such as in physics and engineering problems. In network analysis such as in advanced telecommunication systems studies X traditionally represents the length of service centers in an open/closed queueing network processing jobs. In this context these two quantities (1.1) represent repectively the probability of buffer-overflows and the distribution of the queueing process in this overflow regime.

Several numerical methods have been proposed in the literature to estimate the entrance probability into a rare set. We refer the reader to the excellent paper [9] which contains a precise review on these methods as well as a detailed list of references. For the convenience of the reader we present hereafter a brief description of the two leading ideas.

The first one is based on changing the reference probability so that rare sets becomes less rare. This probabilistic approach often requires the finding of the right change of measure. This step is usually done using large deviations techniques. Another more physical approach consists in splitting the state space into a sequence of sub-levels the particle needs to pass before it reaches the rare target. This splitting stage is based on a precise physical description of the evolution of the process between each level leading to the rare set. The next step is to introduce a system of particles evolving in this level decomposition of the state, in which each particle branches as soon as it enters into a higher level.

The purpose of the present article is to connect the multilevel splitting techniques with the branching and interacting particle systems approximations of Feynman–Kac distributions studied in previous articles. This work has been influenced by the three referenced papers [7, 8] and [9].

Our objective is twofold. First we propose a Feynman–Kac representation of the quantities (1.1). The general idea behind our construction is to consider the level crossing Markov chain in path space and associated with the splitting of the state space. The concatenation of the corresponding states will contain all information on the way the process passes each level before entering into the final and target rare set. Based on this modeling we introduce a natural mean field type genetic particle approximation of the desired quantities (1.1). More interestingly we also show how the genealogical structure of the particle at each level can be used to find the distribution of the process during its excursions to the rare target.

When the state space is splitted into m levels the particle model evolve into m steps. At time $n = 0$ the start with N independent copies of the process X. The particles which enter the recurrent set R are killed and instantly a particle having reached the first level produces an offspring. If the whole system is killed the algorithm is stopped. Otherwise by construction of the birth and death rule we obtain N particles at the first level. At time $n = 1$ the N particles in the first level evolve according to the same rule as the process X. Here again particles which enter the recurrent set R are killed and instantly a particle having reached the second level produces an offspring and so on.

From this brief description we see that the former particle scheme follows the same splitting strategies as the one discussed in [9]. The new mathematical models presented

here allows to calibrate with precision the asymptotic behavior of this particle techniques as the size of the systems tends to infinity. In addition and in contrast to previous articles on the subject the Feynman–Kac analysis in path space presented hereafter allows to study the genealogical structure of these splitting particle models. We will show that the empirical measures associated with the corresponding historical processes converge as $N \to \infty$ to the distribution of the whole path process between each levels.

An empirical method called RESTART [14, 13] can also be used to compute rare transient events and the probability of rare events in steady state, not only the probability to reach the target before coming back to a recurrent set. It was developped to compute the rate of lost packets through a network in a steady–state regime. From a mathematical point of view, this is equivalent to the fraction of time that the trajectory spends in a particular set B, asymptotically as $t \to +\infty$, provided we assume that the system is ergodic. In order to be able to both simulate the system on the long time and see frequent visits to the rare event, the algorithm splits the trajectories crossing the levels "upwards" (getting closer to the rare event), and cancel those crossing "downwards", except one of them called the master trajectory. So the main purpose of this algorithm is quite different from the one of RESTART. It used by practitionners, but this method requires some mathematical approximations which are not yet well understood. Moreover this method is not taken into account by the previous formalism. So, a further work could be an extension of the former particle scheme for covering RESTART.

A short description of the paper is as follows. Section 2 of this paper sets out the Feynman–Kac representation of the quantities (1.1). Section 3 provides the description of the corresponding genetic-type interacting particle system approximating model. Section 4 introduces a path-valued interacting particle systems model for the historical process associated with the previous genetic algorithm. Finally, Section 5 deals with a numerical example, based on the Ornstein-Uhlenbeck process. Estimation of exit time for diffusions controlled by potentials are suggested also, since the lack of exact calculations, even if some heuristics may be applied.

2 Multi-level Feynman–Kac formulae

In practice the process X, before visiting R or entering into the desired set B, passes through a decreasing sequence of closed sets

$$B = B_{m+1} \subset B_m \subset \cdots \subset B_1 \subset B_0 \ .$$

The parameter m and the sequence of level sets depend on the problem at hand. We choose the level sets to be nested to ensure that the process X cannot enter B_n before visiting B_{n-1}. The choice of the recurrent set R depends on the nature of the underlying process X. To visualize these level sets we propose hereafter the two "dual" constructions corresponding to the two "dual" interpretations presented in the introduction.

- In the ballistic regime the decreasing sequence $B = B_{m+1} \subset B_m \subset \cdots \subset B_1 \subset B_0$ represents the physical levels the process X needs to pass before it reaches B.

- In the case of a particle X evolving in a pocket C of the state space S containing "hard obstacles" the sequence $B = B_{m+1} \subset B_m \subset \cdots \subset B_1 \subset B_0$ represents the exit level sets the process needs to reach to get out of C before beeing killed by an obstacle.

To capture the behavior of X between the different levels $B = B_{m+1} \subset B_m \subset \cdots \subset B_1 \subset B_0$ we introduce the discrete event–driven stochastic sequence

$$\mathcal{X}_n = (X_t\,,\, T_{n-1} \wedge T \leq t \leq T_n \wedge T) \in E \quad \text{with} \quad E = \bigcup_{t' \leq t''} \mathbb{D}([t', t''], S)$$

for any $1 \leq n \leq m+1$, where T_n represents the first time X reaches B_n, that is

$$T_n = \inf\{t \geq 0 \ : \ X_t \in B_n\}$$

with the convention $\inf \emptyset = \infty$. At this point we need to endow E with a σ-algebra. First we extend all the trajectories $\mathcal{X}$ by 0 such that they are defined on the whole real line. We denote by $\tilde{\mathcal{X}}$ the corresponding extended trajectory. They are then element of $\mathbb{D}(\mathbb{R}, S)$, on which we consider the σ-algebra generated by the Skorohod metric. Then we consider the product space $\tilde{E} = \mathbb{D}(\mathbb{R}, S) \times \bar{\mathbb{R}}_+ \times \bar{\mathbb{R}}_+$ endowed with the product σ-algebra. Finally to any element $\mathcal{X} \in E$, defined on an interval $[s, t]$, we associate $(\tilde{\mathcal{X}}, s, t) \in \tilde{E}$. So we have imbedded E in $\tilde{E}$ in such a way that all the standard functionals $(\sup, \inf, \dots)$ have good measurability properties. We denote by $\mathcal{B}_b(E)$ the measurable bounded functions from E (or equivalently its image in $\tilde{E}$) into $\mathbb{R}$.

Notice that

- if $T < T_{n-1}$, then $\mathcal{X}_n = \{X_T\}$ and $X_{T_n \wedge T} = X_T \notin B_n$,
- if $T_{n-1} \leq T < T_n$, then $\mathcal{X}_n = (X_t\,,\, T_{n-1} \leq t \leq T)$ and $X_{T_n \wedge T} = X_T \notin B_n$,
- finally, if $T_n \leq T$, then $\mathcal{X}_n = (X_t\,,\, T_{n-1} \leq t \leq T_n)$ represents the path of X between the successive levels B_{n-1} and B_n, and $X_{T_n \wedge T} = X_{T_n} \in B_n$.

Consequently, $X_{T_n \wedge T} \in B_n$ if and only if $T_n \leq T$. By construction we also notice that

$$T_0 = 0 \leq T_1 \leq \cdots \leq T_m \leq T_{m+1} = T_B$$

and for each n

$$(T_{n-1} > T) \Rightarrow (T_p > T \text{ and } \mathcal{X}_p = \{X_T\} \not\subset B_p, \text{ for all } p \geq n)\ .$$

From these observation we can alternatively define the times T_n by the inductive formula

$$T_n = \inf\{t \geq T_{n-1} \ : \ X_t \in B_n\}$$

with the convention $\inf \emptyset = \infty$ so that $T_n > T$ if either $T_{n-1} > T$ or if starting in B_{n-1} at time T_{n-1} the process never reaches B_n before time T. We also observe that

$$(T_B \leq T) \Leftrightarrow (T_{m+1} \leq T) \Leftrightarrow (T_1 \leq T, \cdots, T_{m+1} \leq T)$$

By the strong Markov property we check that the stochastic sequence $(\mathcal{X}_0\,,\cdots,\mathcal{X}_{m+1})$ forms an E-valued Markov chain. One way to check whether the path has succeeded to reach the desired n-th level is to consider the potential functions g_n on E defined for each $x=(x_t\,,\,t'\leq t\leq t'')\in\mathbb{D}([t',t''],S)$ with $t'\leq t''$ by

$$g_n(x)=1_{(x_{t''}\,\in\,B_n)}$$

In this notation, we have for each n

$$(T_n\leq T)\Leftrightarrow(T_1\leq T,\cdots,T_n\leq T)\Leftrightarrow(g_1(\mathcal{X}_1)=1,\cdots,g_n(\mathcal{X}_n)=1)\;,$$

i.e.

$$1_{(T_n\,\leq\,T)}=\prod_{k=0}^{n}g_k(\mathcal{X}_k)\;,$$

and

$$f(\mathcal{X}_n)\;1_{(T_n\,\leq\,T)}=f(X_t\,,\,T_{n-1}\leq t\leq T_n)\;1_{(T_n\,\leq\,T)}\;.$$

For later purpose, we introduce the following notation

$$\begin{aligned}(\mathcal{X}_0,\cdots,\mathcal{X}_n)&=(X_0,(X_t\,,\,0\leq t\leq T_1\wedge T),\cdots,(X_t\,,\,T_{n-1}\wedge T\leq t\leq T_n\wedge T))\\&=[X_t\,,\,0\leq t\leq T_n\wedge T]\;.\end{aligned}$$

Introducing the Feynman–Kac distribution η_n defined by

$$\eta_n(f)=\frac{\gamma_n(f)}{\gamma_n(1)}\quad\text{with}\quad\gamma_n(f)=\mathbb{E}(f(\mathcal{X}_n)\prod_{k=0}^{n}g_k(\mathcal{X}_k))\;,$$

for any bounded measurable function f defined on E, we are now able to state the following Feynman–Kac representation of the quantities (1.1).

Theorem 1 (Multilevel Feynman–Kac formula) *For any n and for any $f\in\mathcal{B}_b(E)$ we have*

$$\mathbb{E}(f(X_t\,,\,T_{n-1}\leq t\leq T_n)\mid T_n\leq T)=\frac{\mathbb{E}(f(\mathcal{X}_n)\prod_{p=0}^{n}g_p(\mathcal{X}_p))}{\mathbb{E}(\prod_{p=0}^{n}g_p(\mathcal{X}_p))}=\frac{\gamma_n(f)}{\gamma_n(1)}=\eta_n(f)\;,$$

and

$$\mathbb{P}(T_n\leq T)=\mathbb{E}(\prod_{k=0}^{n}g_k(\mathcal{X}_k))=\gamma_n(1)\;.$$

In addition for any $f \in \mathcal{B}_b(E^{n+1})$ *we have that*

$$\mathbb{E}(f([X_t\,,\,0\leq t\leq T_n])\mid T_n\leq T)=\frac{\mathbb{E}(f(\mathcal{X}_0,\cdots,\mathcal{X}_n)\prod_{p=0}^{n}g_p(\mathcal{X}_p))}{\mathbb{E}(\prod_{p=0}^{n}g_p(\mathcal{X}_p))}$$

The straightforward formula

$$\mathbb{P}[T_n\leq T]=\prod_{k=0}^{n}\mathbb{P}[T_k\leq T\mid T_{k-1}\leq T]\;,$$

which shows how the very small probability of a rare event can be decomposed into the product of reasonably small but not too small conditional probabilities, each of which corresponding to the transition between two events, can be recoverd from the the well–known identity

$$\gamma_n(1)=\prod_{k=0}^{n}\eta_{k-1}(g_k)\;,$$

and will provide the basis for the efficient numerical approximation in terms of an interacting particle system. These conditional probabilities are not known in advance, and are learned by the algorithm as well.

3 Genetic approximating models

In previous studies [7, 8] we design a collection of branching and interacting particle systems approximating models for solving a general class of Feynman–Kac models. These particle techniques can be used to solve the formulae presented in Theorem 1.1. We first focus on a simple mutation/selection genetic algorithm.

3.1 Classical scheme

To describe this particle approximating model we first recall that the Feynman–Kac distribution flow $\eta_n\in\mathcal{P}(E)$ defined by

$$\eta_n(f)=\frac{\gamma_n(f)}{\gamma_n(1)}\quad\text{with}\quad\gamma_n(f)=\mathbb{E}(f(\mathcal{X}_n)\prod_{p=0}^{n-1}g_p(\mathcal{X}_p))$$

is solution of the following measure valued dynamical system

$$\eta_{n+1}=\Phi_{n+1}(\eta_n)\tag{3.1}$$

The mappings Φ_{n+1} from the set of measures

$$\mathcal{P}_n(E) = \{\eta \in \mathcal{P}(E)\,,\, \eta(g_n) > 0\}$$

into $\mathcal{P}(E)$ are defined by

$$\Phi_{n+1}(\eta)(dx') = (\Psi_n(\eta)\,\mathcal{K}_{n+1})(dx') = \int_E \Psi_n(\eta)(dx)\,\mathcal{K}_{n+1}(x, dx')$$

The Markov kernels $\mathcal{K}_n(x, dx')$ represent the Markov transitions of the chain $\mathcal{X}_n$. The updating mappings Ψ_n are defined from $\mathcal{P}_n(E)$ into $\mathcal{P}_n(E)$ and for any $\eta \in \mathcal{P}_n(E)$ and $f \in \mathcal{B}_b(E)$ by the formula

$$\Psi_n(\eta)(f) = \eta(f\,g_n)/\eta(g_n)$$

Thus we see that the recursion (3.1) involves two separate selection / mutation transitions

$$\eta_n \in \mathcal{P}(E) \xrightarrow{\text{selection}} \widehat{\eta}_n = \Psi_n(\eta_n) \in \mathcal{P}(E) \xrightarrow{\text{mutation}} \eta_{n+1} = \widehat{\eta}_n\,\mathcal{K}_{n+1} \in \mathcal{P}(E) \tag{3.2}$$

It is also conventient to recall that the finite and positive measures γ_n on E can be expressed in terms of the flow $\{\eta_0, \cdots, \eta_n\}$, using the easily checked formula

$$\gamma_n(f) = \eta_n(f) \prod_{p=0}^{n-1} \eta_p(g_p)$$

In these notations we readily observe that

$$\gamma_n(g_n) = \mathbb{P}(T_n \leq T)$$

and

$$\widehat{\eta}_n(f) = \Psi_n(\eta_n)(f) = \mathbb{E}(f(X_t\,,\, T_{n-1} \leq t \leq T_n) \mid T_n \leq T)$$

The genetic type N–particle system associated with an abstract measure valued process of the form (3.1) is the Markov chain $\xi_n = (\xi_n^1, \cdots, \xi_n^N)$ taking values at each time n in the product state spaces $E^N \cup \{\Delta\}$ where Δ stands for a cemetery or coffin point. Its transitions are defined as follows. For any configuration $x = (x^1, \cdots, x^N) \in E^N$ such that $\frac{1}{N}\sum_{i=1}^N \delta_{x^i} \in \mathcal{P}_n(E)$ we set

$$\mathbb{P}(\xi_{n+1} \in dy \mid \xi_n = x) = \prod_{p=1}^N \Phi_{n+1}(\frac{1}{N}\sum_{i=1}^N \delta_{x^i})(dy^p) \tag{3.3}$$

where $dy = dy^1 \times \cdots \times dy^N$ is an infinitesimal neighborhood of $y = (y^1, \cdots, y^N) \in E^N$. When the system arrives in some configuration $\xi_n = x$ such that

$$\frac{1}{N}\sum_{i=1}^{N} \delta_{x^i} \notin \mathcal{P}_n(E)$$

the particle algorithm is stopped and we set $\xi_{n+1} = \Delta$. The initial system of particles $\xi_0 = (\xi_0^1, \cdots, \xi_0^N)$ consists in N independent random variables with common law $\eta_0 = \mathrm{Law}(\mathcal{X}_0) = \mathrm{Law}(X_0)$. The superscript $i = 1, \cdots, N$ represents the label of the particle and the parameter N is the size of the systems and the precision of the algorithm.

Next we describe in more details the genetic evolution of the path-particles. At the time $n = 0$ the initial configuration consists in N independent and identically distributed S-valued random variables ξ_0^i with common law η_0. Since we have $g_0(u) = 1$ for η_0-almost every $u \in S$, we may discard the selection at time $n = 0$ and set $\widehat{\xi}_0^i = \xi_0^i$ for each $1 \leq i \leq N$. If we use the convention $T_{-1}^i = \widehat{T}_{-1}^i = 0$ and if we set $T_0^i = \widehat{T}_0^i = 0$ we notice that the single states ξ_0^i and $\widehat{\xi}_0^i$ can be written in the path-form

$$\xi_0^i = \xi_0^i(0) = (\xi_0^i(t)\,,\, T_{-1}^i \leq t \leq T_0^i) \quad \text{and} \quad \widehat{\xi}_0^i = \widehat{\xi}_0^i(0) = (\widehat{\xi}_0^i(t)\,,\, \widehat{T}_{-1}^i \leq t \leq \widehat{T}_0^i)$$

The **mutation transition** $\widehat{\xi}_n \to \xi_{n+1}$ at time $(n+1)$ is defined as follows. If $\widehat{\xi}_n = \Delta$ we set $\xi_{n+1} = \Delta$. Otherwise during mutation, independently of each other, each selected path-particle

$$\widehat{\xi}_n^i = (\widehat{\xi}_n^i(t)\,,\, \widehat{T}_n^{-,i} \leq t \leq \widehat{T}_n^{+,i})$$

evolves randomly according to the Markov transition $\mathcal{K}_{n+1}$ of the Markov chain $\mathcal{X}_{n+1}$ at time $(n+1)$ so that

$$\xi_{n+1}^i = (\xi_{n+1}^i(t)\,,\, T_{n+1}^{-,i} \leq t \leq T_{n+1}^{+,i})$$

is a random variable with distribution $\mathcal{K}_{n+1}(\widehat{\xi}_n^i, dx')$.

In other words, the algorithm goes like this between steps n and $n+1$. For each particle i we start a trajectory from $\widehat{\xi}_n^i$ at time $T_{n+1}^{-,i} = \widehat{T}_n^{+,i}$, and let it evolve randomly as a copy $\{\xi_{n+1}^i(s)\,,\, s \geq T_{n+1}^{-,i}\}$ of the process $\{X_s\,,\, s \geq T_{n+1}^{-,i}\}$, until the stopping time $T_{+,n+1}^i$, which is either

$$T_{n+1}^{+,i} = \inf\{t \geq T_{n+1}^{-,i} \;:\; \xi_{n+1}^i(t) \in B_{n+1} \cup R\},$$

in case of a recurrent set to be avoided, or

$$T_{n+1}^{+,i} = T \wedge \inf\{t \geq T_{n+1}^{-,i} \;:\; \xi_{n+1}^i(t) \in B_{n+1}\},$$

in case of a deterministic final time, depending on the problem at hand.

The **selection transition** $\xi_{n+1} \rightarrow \widehat{\xi}_{n+1}$ is defined as follows. From the previous mutation transition we obtain N path-particle

$$\xi^i_{n+1} = (\xi^i_{n+1}(t)\,,\, T^{-,i}_{n+1} \leq t \leq T^{+,i}_{n+1})$$

Only some of these particle have succeeded to reach to desired set B_{n+1} and the other ones have failed. We denote by I^N_{n+1} the labels of the particles having succeeded to reach the $(n+1)$-th level

$$I^N_{n+1} = \{i = 1, \cdots, N \,:\, \xi^i_{n+1}(T^{+,i}_{n+1}) \in B_{n+1}\}$$

If $I^N_{n+1} = \emptyset$ then none of the particles have succeeded to reach the desired level. Since

$$I^N_{n+1} = \emptyset \iff \frac{1}{N}\sum_{i=1}^N g_{n+1}(\xi^i_{n+1}) = 0 \iff \frac{1}{N}\sum_{i=1}^N \delta_{\xi^i_{n+1}} \notin \mathcal{P}_{n+1}(E)$$

we see that in this situation the algorithm is stopped and $\widehat{\xi}_{n+1} = \Delta$. Otherwise the selection transition of the N-particle models (3.3) and (3.5) are defined as follows. In the first situation the system $\widehat{\xi}_{n+1} = (\widehat{\xi}^1_{n+1}, \cdots, \widehat{\xi}^N_{n+1})$ consists in N independent (given the past until the last mutation) random variables

$$\widehat{\xi}^i_{n+1} = (\widehat{\xi}^i_{n+1}(t)\,,\, \widehat{T}^{-,i}_{n+1} \leq t \leq \widehat{T}^{+,i}_{n+1})$$

with common distribution

$$\Psi_{n+1}(\frac{1}{N}\sum_{i=1}^N \delta_{\xi^i_{n+1}}) = \sum_{i=1}^N \frac{g_{n+1}(\xi^i_{n+1})}{\sum_{j=1}^N g_{n+1}(\xi^j_{n+1})}\,\delta_{\xi^i_{n+1}} = \frac{1}{|I^N_{n+1}|}\sum_{i\in I^N_{n+1}} \delta_{(\xi^i_{n+1}(t)\,,\, T^{-,i}_{n+1} \leq t \leq T^{+,i}_{n+1})}$$

In simple words, we draw them uniformly among the sucessfull pieces of trajectories $\{\xi^i_{n+1}, i \in I^N_{n+1}\}$.

3.2 Alternate scheme

As mentioned above the choice of the N-particle approximating model of (3.1) is not unique. Below, we propose an alternative scheme which contains in some sense less randomness [6]. The key idea is to notice that the updating mapping $\Psi_n : \mathcal{P}_n(E) \rightarrow \mathcal{P}_n(E)$ can be rewritten in the following form

$$\Psi_n(\eta)(dx') = (\eta\, \mathcal{S}_n(\eta))(dx') = \int_E \eta(dx)\, \mathcal{S}_n(\eta)(x, dx')\ , \tag{3.4}$$

with the collection of Markov transition kernels $\mathcal{S}_n(\eta)(x, dx')$ on E defined by

$$\mathcal{S}_n(\eta)(x, dx') = (1 - g_n(x))\ \Psi_n(\eta)(dx') + g_n(x)\ \delta_x(dx')\ ,$$

where

$$g_n(x) = 1_{(g_n(x)=1)} = 1_{(x \in g_n^{-1}(1))} ,$$

and where $g_n^{-1}(1)$ stands for the set of paths in E entering the level B_n, that is

$$g_n^{-1}(1) = \{x \in E : g_n(x) = 1\} = \{x \in \mathbb{D}([t', t''], S) , t' \leq t'' : x_{t''} \in B_n\} .$$

Indeed

$$(\eta \, \mathcal{S}_n(\eta))(dx') = \Psi_n(\eta)(dx') \, (1 - \eta(g_n)) + \int_E \eta(dx) \, g_n(x) \, \delta_x(dx') ,$$

hence

$$(\eta \, \mathcal{S}_n(\eta))(f) = \Psi_n(\eta)(f) \, (1 - \eta(g_n)) + \eta(f \, g_n) = \Psi_n(\eta)(f) ,$$

for any bounded measurable function f defined on E, which proves (3.4). In this notation, (3.1) can be rewritten as

$$\eta_{n+1} = \eta_n \, \mathcal{K}_{n+1}(\eta_n) ,$$

with the composite Markov transition kernel $\mathcal{K}_{n+1}(\eta)$ defined by

$$\mathcal{K}_{n+1}(\eta)(x, dx') = (\mathcal{S}_n(\eta) \, \mathcal{K}_{n+1})(x, dx') = \int_E \mathcal{S}_n(\eta)(x, dx'') \, \mathcal{K}_{n+1}(x'', dx')$$

The alternative N-particle model associated with this new description is defined as before by replacing (3.3) by

$$\mathbb{P}(\xi_{n+1} \in dy \mid \xi_n = x) = \prod_{p=1}^{N} \mathcal{K}_{n+1}(\frac{1}{N} \sum_{i=1}^{N} \delta_{x^i})(x^p, dy^p) \tag{3.5}$$

By definition of Φ_{n+1} and $\mathcal{K}_{n+1}(\eta)$ we have for any configuration $x = (x^1, \cdots, x^N) \in E^N$ with $\frac{1}{N} \sum_{i=1}^{N} \delta_{x^i} \in \mathcal{P}_n(E)$

$$\Phi_{n+1}(\frac{1}{N} \sum_{i=1}^{N} \delta_{x^i})(dv) = \sum_{i=1}^{N} \frac{g_n(x^i)}{\sum_{j=1}^{N} g_n(x^j)} \, \mathcal{K}_{n+1}(x^i, dv)$$

In much the same way we find that

$$\mathcal{K}_{n+1}(\frac{1}{N} \sum_{i=1}^{N} \delta_{x^i}) = \mathcal{S}_n(\frac{1}{N} \sum_{i=1}^{N} \delta_{x^i}) \, \mathcal{K}_{n+1}$$

with the selection transition

$$\mathcal{S}_n(\frac{1}{N}\sum_{i=1}^{N}\delta_{x^i})(x^p,dv) = (1-g_n(x^p))\ \Psi_n(\frac{1}{N}\sum_{i=1}^{N}\delta_{x^i})(dv) + g_n(x^p)\,\delta_{x^p}(dv)$$

where

$$\Psi_n(\frac{1}{N}\sum_{i=1}^{N}\delta_{x^i}) = \sum_{i=1}^{N}\frac{g_n(x^i)}{\sum_{j=1}^{N} g_n(x^j)}\,\delta_{x^i}$$

Thus, we see that the transition $\xi_n \to \xi_{n+1}$ of the former Markov models splits up into two separate genetic type mechanisms

$$\xi_n \in E^N \cup \{\Delta\} \xrightarrow{\text{selection}} \widehat{\xi}_n = (\widehat{\xi}_n^i)_{1\leq i\leq N} \in E^N \cup \{\Delta\} \xrightarrow{\text{mutation}} \xi_{n+1} \in E^N \cup \{\Delta\}$$

By construction we notice that

$$\xi_n = \Delta \implies \forall p \geq n \quad \xi_p = \Delta \quad \text{and} \quad \widehat{\xi}_p = \Delta$$

By definition of the path valued Markov chain $\mathcal{X}_n$ this genetic model consists in N-path valued particles

$$\begin{aligned}\xi_n^i &= (\xi_n^i(t)\,,\, T_n^{-,i} \leq t \leq T_n^{+,i}) \in \mathbb{D}([T_n^{-,i}, T_n^{+,i}], S)\\ \widehat{\xi}_n^i &= (\widehat{\xi}_n^i(t)\,,\, \widehat{T}_n^{-,i} \leq t \leq \widehat{T}_n^{+,i}) \in \mathbb{D}([\widehat{T}_n^{-,i}, \widehat{T}_n^{+,i}], S)\ .\end{aligned}$$

The random time-pairs $(T_n^{-,i}, T_n^{+,i})$ and $(\widehat{T}_n^{-,i}, \widehat{T}_n^{+,i})$ represent the first and last time of the corresponding paths.

In the alternative model (3.5) each particle

$$\widehat{\xi}_{n+1}^i = (\widehat{\xi}_{n+1}^i(t)\,,\, \widehat{T}_{n+1}^{-,i} \leq t \leq \widehat{T}_{n+1}^{+,i})$$

is sampled according to the selection distribution

$$\begin{aligned}&\mathcal{S}_{n+1}(\frac{1}{N}\sum_{j=1}^{N}\delta_{\xi_{n+1}^j})(\xi_{n+1}^i, dv)\\ &= (1-g_{n+1}(\xi_{n+1}^i))\ \Psi_n(\frac{1}{N}\sum_{j=1}^{N}\delta_{\xi_{n+1}^j})(dv) + g_{n+1}(\xi_{n+1}^i)\ \delta_{\xi_{n+1}^i}(dv)\\ &= 1_{(\xi_{n+1}^i(T_{n+1}^{+,i}) \notin B_{n+1})}\ \Psi_n(\frac{1}{N}\sum_{j=1}^{N}\delta_{\xi_{n+1}^j})(dv) + 1_{(\xi_{n+1}^i(T_{n+1}^{+,i}) \in B_{n+1})}\ \delta_{\xi_{n+1}^i}(dv)\end{aligned}$$

More precisely we have

$$\xi_{n+1}^i(T_{n+1}^{+,i}) \in B_{n+1} \implies \widehat{\xi}_{n+1}^i = \xi_{n+1}^i \ .$$

In the opposite we have $\xi_{n+1}^i(T_{n+1}^{+,i}) \notin B_{n+1}$ when the particle has not succeeded to reach the $(n+1)$–th level. In this case $\widehat{\xi}_{n+1}^i$ is chosen randomly and uniformly in the set

$$\{\xi_{n+1}^j \ : \ \xi_{n+1}^j(T_{n+1}^{+,j}) \in B_{n+1}\} = \{\xi_{n+1}^j \ : \ j \in I_{n+1}^N\} \ ,$$

of all particle having succeeded to enter into B_{n+1}. In other words each particle which does not enter into the $(n+1)$–th level is killed and instantly a different particle in the B_{n+1} level splits into two offsprings.

We denote by τ^N the lifetime of the N-genetic model

$$\tau^N = \inf\{n \geq 0 \ : \ \frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i} \notin \mathcal{P}_n(E)\} \ .$$

For each time $n < \tau^N$ we denote by η_n^N and $\widehat{\eta}_n^N$ the particle density profiles associated with the N-particle model

$$\eta_n^N = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i} \quad \text{and} \quad \widehat{\eta}_n^N = \Psi_n(\eta_n^N) \ .$$

For each time $n < \tau^N$ the N–particle approximating measures γ_n^N associated with γ_n are defined for any $f \in \mathcal{B}_b(E)$ by

$$\gamma_n^N(f) = \eta_n^N(f) \prod_{p=0}^{n-1} \eta_p^N(g_p) \ .$$

Note that

$$\gamma_n^N(g_n) = \prod_{p=0}^n \eta_p^N(g_p) = \prod_{p=1}^n \frac{|I_p^N|}{N} \ ,$$

and

$$\widehat{\eta}_n^N = \Psi_n(\eta_n^N) = \frac{1}{|I_n^N|} \sum_{i \in I_n^N} \delta_{(\xi_n^i(t) \, , \, T_n^{-,i} \leq t \leq T_n^{+,i})} \ .$$

The asymptotic behavior as $N \to \infty$ of the interacting particle model we have constructed has been studied in many works. We refer the reader to the survey paper [7] in the case of strictly positive potentials g_n and [4, 5] for non negative potentials. For the convenience of the reader we have chosen to present the impact of some exponential and $\mathbb{L}_p$-mean error estimates, and a fluctuation result, in the analysis of rare events.

Theorem 2 *For any $0 \leq n \leq m+1$ there exists a finite constant c_n such that for any $N \geq 1$*

$$\mathbb{P}(\tau^N \leq n) \leq c_n \exp\left(-N/c_n\right) .$$

The particle estimates $\gamma_n^N(g_n)$ are unbiased

$$\mathbb{E}(\gamma_n^N(g_n)\, 1_{(\tau^N > n)}) = \mathbb{P}(T_n \leq T)$$

and for each $p \geq 1$ we have

$$(\mathbb{E}|\, \gamma_n^N(g_n)\, 1_{(\tau^N > n)} - \mathbb{P}(T_n \leq T)\,|^p)^{1/p} \leq a_p\, b_n/\sqrt{N} ,$$

for some finite constant $a_p < \infty$ which only depends on the parameter p, and for some finite constant $b_n < \infty$ which only depends on the time parameter n. In addition, for any test function $f \in \mathcal{B}_b(E)$, with $\|f\| \leq 1$

$$(\mathbb{E}|\, \widehat{\eta}_n^N(f)\, 1_{(\tau^N > n)} - \mathbb{E}(f(X_t\,, T_{n-1} \leq t \leq T_n) \mid T_n \leq T)\,|^p)^{1/p} \leq a_p\, b_n/\sqrt{N} .$$

We illustrate the impact of this asymptotic convergence theorem by chosing some particular test functions. For each $u > 0$ we define the function $f^{(u)}$ on E by setting for each $x = (x_r\,,\, s \leq r \leq t) \in \mathbb{D}([s,t], S)$ with $s \leq t$,

$$f^{(u)}(x) = \left\{ \begin{array}{lll} 1 & \text{if} & |t-s| \leq u \\ 0 & \text{if} & |t-s| > u \end{array} \right. \tag{3.6}$$

In this notation $u \to \Psi_n(\eta_n)(f^{(u)})$ is the repartition function of the intertime $T_n - T_{n-1}$ between two consecutive levels B_{n-1} and B_n, that is

$$\Psi_n(\eta_n)(f^{(u)}) = \mathbb{P}(T_n - T_{n-1} \leq u \mid T_n \leq R)$$

The particle approximation of this quantity is the proportion of paths having passed from B_{n-1} to B_n in time u.

Now a CLT-type result on the error fluctuations. Let us first introduce the following notation:

$$a_n = \sum_{p=0}^{n} \mathbb{E}\left[[\Delta_{p-1,p}^n(T_p, X_{T_p})\, 1_{T_p \leq T} - 1]^2 \,|T_{p-1} \leq T\right],$$

and

$$b_n = \sum_{p=0}^{n} \mathbb{E}\left[1_{T_p \leq T}\, [\Delta_{p,p}^n(T_p, X_{T_p}) - 1]^2 \,|T_{p-1} \leq T\right],$$

with the functions $\Delta_{p,q}^n$ defined by

$$\Delta_{p,q}^n(t,x) = \frac{\mathbb{P}(T_n \leq T | T_q = t, X_{T_q} = x)}{\mathbb{P}(T_n \leq T | T_p \leq T)}.$$

Theorem 3 *For any $0 \leq n \leq m$, the sequence of random variables*

$$W_{n+1}^N = \sqrt{N}\left(1_{\tau^N<n}\,\gamma_{n+1}^N(1) - \mathbb{P}(T_n \leq T)\right)$$

converges in distribution to a Gaussian $N(0,\sigma_n^2)$, with

$$\sigma_n^2 = \mathbb{P}(T_n \leq T)^2(a_n - b_n).$$

We end this section with a physical discussion on the terms a_n, and b_n.

Proposition 4 *For any time horizon, we have the formula*

$$\begin{aligned} a_n - b_n &= \sum_{p=0}^{n}\left(\frac{1}{\mathbb{P}(T_p \leq T | T_{p-1} \leq T)} - 1\right) \\ &\quad + \sum_{p=0}^{n}\mathbb{E}\left[\left[\frac{\mathbb{P}(T_n \leq T | T_p, X_{T_p})}{\mathbb{P}(T_n \leq T | T_p \leq T)} - 1\right]^2 \; | T_p \leq T\right] \\ &\qquad \times \left[\frac{1}{\mathbb{P}(T_p \leq T | T_{p-1} \leq T)} \; - \mathbb{P}(T_p \leq T | T_{p-1} \leq T)\right]. \end{aligned}$$

Proof:
Firstly, we observe that

$$\begin{aligned} \mathbb{E}\left[\Delta_{p-1,p}^n(T_p, X_{T_p}) \mid T_p \leq T\right] &= \mathbb{E}\left[\frac{\mathbb{P}(T_n \leq T | T_p, X_{T_p})}{\mathbb{P}(T_n \leq T | T_{p-1} \leq T)} \mid T_p \leq T\right] \\ &= \frac{\mathbb{P}(T_n \leq T | T_p \leq T)}{\mathbb{P}(T_n \leq T | T_{p-1} \leq T)}. \end{aligned}$$

Using the fact that

$$q \geq p \Longrightarrow \mathbb{P}(T_q \leq T \;,\; T_p \leq T) = \mathbb{P}(T_q \leq T)$$

we conclude that

$$\mathbb{E}\left[\Delta_{p-1,p}^n(T_p, X_{T_p}) \mid T_p \leq T\right] = \frac{1}{\mathbb{P}(T_p \leq T | T_{p-1} \leq T)}. \tag{3.7}$$

In much the same way, we observe that

$$\frac{\mathbb{E}\left[f_p(T_p, X_{T_p})\; 1_{T_p \leq T} \mid T_{p-1} \leq T\right]}{\mathbb{E}\left[\; 1_{T_p \leq T} \mid T_{p-1} \leq T\right]} = \mathbb{E}\left[f_p(T_p, X_{T_p}) \mid T_p \leq T\right]$$

for any measurable function f_p on $(\mathbb{R}_+ \times S)$. This yields that

$$\begin{aligned} &\mathbb{E}\left[f_p(T_p, X_{T_p})\; 1_{T_p \leq T} \mid T_{p-1} \leq T\right] \\ &= \mathbb{E}\left[f_p(T_p, X_{T_p}) \mid T_p \leq T\right] \times \mathbb{P}(T_p \leq T | T_{p-1} \leq T). \end{aligned}$$

Using (3.7), we find that

$$\begin{aligned}
&\mathbb{E}\left[\Delta^n_{p-1,p}(T_p, X_{T_p})\ 1_{T_p \le T} | T_{p-1} \le T\right] \\
&= \mathbb{E}\left[\Delta^n_{p-1,p}(T_p, X_{T_p}) | T_p \le T\right] \times \mathbb{P}(T_p \le T | T_{p-1} \le T) = 1.
\end{aligned}$$

From the above observations, we arrive at

$$\begin{aligned}
&\mathbb{E}\left[[\Delta^n_{p-1,p}(T_p, X_{T_p})\ 1_{T_p \le T} - 1]^2\ | T_{p-1} \le T\right] \\
&= \mathbb{E}\left[[\Delta^n_{p-1,p}(T_p, X_{T_p})]^2\ | T_p \le T\right] \times \mathbb{P}(T_p \le T | T_{p-1} \le T) - 1.
\end{aligned}$$

Using again (3.7), we end up with the following formula

$$\begin{aligned}
&\mathbb{E}\left[[\Delta^n_{p-1,p}(T_p, X_{T_p})\ 1_{T_p \le T} - 1]^2\ | T_{p-1} \le T\right] \\
&= \mathbb{E}\left[\left[\frac{\Delta^n_{p-1,p}(T_p, X_{T_p})}{\mathbb{E}\left[\Delta^n_{p-1,p}(T_p, X_{T_p})\ | T_p \le T\right]}\right]^2\ | T_p \le T\right] \times \frac{1}{\mathbb{P}(T_p \le T | T_{p-1} \le T)} - 1.
\end{aligned}$$

Next, we see that

$$\begin{aligned}
&\mathbb{E}\left[[\Delta^n_{p-1,p}(T_p, X_{T_p})\ 1_{T_p \le T} - 1]^2\ | T_{p-1} \le T\right] \\
&= \left(\frac{1}{\mathbb{P}(T_p \le T | T_{p-1} \le T)} - 1\right) \\
&\quad + \mathbb{E}\left[\left[\frac{\Delta^n_{p-1,p}(T_p, X_{T_p})}{\mathbb{E}\left[\Delta^n_{p-1,p}(T_p, X_{T_p})\ | T_p \le T\right]} - 1\right]^2\ | T_p \le T\right] \times \frac{1}{\mathbb{P}(T_p \le T | T_{p-1} \le T)}
\end{aligned}$$

$$\begin{aligned}
a_n &= \sum_{p=0}^{n} \left(\frac{1}{\mathbb{P}(T_p \le T | T_{p-1} \le T)} - 1\right) \\
&\quad + \sum_{p=0}^{n} \frac{1}{\mathbb{P}(T_p \le T | T_{p-1} \le T)}\ \mathbb{E}\left[\left[\frac{\Delta^n_{p-1,p}(T_p, X_{T_p})}{\mathbb{E}\left[\Delta^n_{p-1,p}(T_p, X_{T_p})\ | T_p \le T\right]} - 1\right]^2\ | T_p \le T\right].
\end{aligned}$$

To take the final step, we observe that

$$\begin{aligned}
\frac{\Delta^n_{p-1,p}(T_p, X_{T_p})}{\mathbb{E}\left[\Delta^n_{p-1,p}(T_p, X_{T_p})\ | T_p \le T\right]} &= \frac{\mathbb{P}(T_n \le T | T_p, X_{T_p})}{\mathbb{E}\left[\mathbb{P}(T_n \le T | T_p, X_{T_p})\ | T_p \le T\right]} \\
&= \frac{\mathbb{P}(T_n \le T | T_p, X_{T_p})}{\mathbb{P}(T_n \le T | T_p \le T)} = \Delta^n_{p,p}(T_p, X_{T_p}).
\end{aligned}$$

This ends the proof of the proposition. ∎

Now we explain the meaning of this proposition. If $\mathbb{P}(T_n \leq T | T_p, X_{T_p})$ does not depend on (T_p, X_{T_p}) given $(T_n \leq T)$, i.e. does not depend on the hitting time and point of the level set B_p, then

$$\mathbb{E}\left[\left[\frac{\mathbb{P}(T_n \leq T | T_p, X_{T_p})}{\mathbb{P}(T_n \leq T | T_p \leq T)} - 1\right]^2 \; | T_p \leq T\right] = 0$$

and if this holds for any $p = 0, 1, \ldots, n$, then the asymptotic variance reduces to the expression

$$\sigma_n^2 = \sum_{p=0}^{n} \left(\frac{1}{\mathbb{P}(T_p \leq T | T_{p-1} \leq T)} - 1\right)$$

as given in [10]. Idealy, the level set B_p should be chosen such that $\mathbb{P}(T_n \leq T | T_p, X_{T_p})$ does not depend on (T_p, X_{T_p}) given $(T_n \leq T)$. Even if this is clearly unrealistic for most practical problems, this observation gives an insight on how to choose the level sets.

4 Genealogical tree based models

The genetic particle approximating model described in the previous section can be interpretated as a birth and death particle model. The particle dies if it does not succeed to reach the desired level and it duplicates in some offsprings when it hits this level. One way to model the genealogical tree and the line of ancestors of the particles alive at some given date is to consider the stochastic sequence

$$\mathcal{Y}_n = (\mathcal{X}_0, \cdots, \mathcal{X}_n) \in E_n = \underbrace{E \times \cdots \times E}_{(n+1)\text{–times}}$$

It is not difficult to check that $\mathcal{Y}_n$ forms a time inhomogenous Markov chain with Markov transitions $\mathcal{Q}_{n+1}$ from E_n into E_{n+1}

$$\mathcal{Q}_{n+1}(x_0, \cdots, x_n, dx'_0, \cdots, dx'_n, dx'_{n+1}) = \delta_{(x_0, \cdots, x_n)}(dx'_0, \cdots, dx'_n)\, \mathcal{K}_{n+1}(x'_n, dx'_{n+1})$$

Let h_n be the mapping from E_n into $[0, \infty)$ defined by

$$h_n(x_0, \cdots, x_n) = g_n(x_n)$$

In this notation we have for any $f_n \in \mathcal{B}_b(E_n)$ the Feynman–Kac representation

$$\begin{aligned}\widehat{\mu}_n(f_n) &= \frac{\mathbb{E}(f_n(\mathcal{Y}_n)\prod_{p=0}^{n} h_p(\mathcal{Y}_p))}{\mathbb{E}(\prod_{p=0}^{n} h_p(\mathcal{Y}_p))} \\ &= \mathbb{E}(f_n(X_0,(X_t\,,\,0\leq t\leq T_1),\cdots,(X_t\,,\,T_{n-1}\leq t\leq T_n))\mid T_n\leq T) \\ &= \mathbb{E}(f_n([X_t\,,\,0\leq t\leq T_n])\mid T_n\leq T)\end{aligned}$$

Using the same lines of reasoning as above the N-particle approximating model associated with these Feynman–Kac distributions is again a genetic algorithm with mutation transitions $\mathcal{Q}_n$ and potential functions h_n. Here the path-particle at time n take values in E_n and they can be written as follows

$$\zeta_n^i=(\xi_{0,n}^i,\cdots,\xi_{n,n}^i)\quad\text{and}\quad\widehat{\zeta}_n^i=(\widehat{\xi}_{0,n}^i,\cdots,\widehat{\xi}_{n,n}^i)\in E_n$$

with for each $0\leq p\leq n$

$$\xi_{p,n}^i=(\xi_{p,n}^i(t)\,,\,T_{p-1,n}^i\leq t\leq T_{p,n}^i)\quad\text{and}\quad\widehat{\xi}_{p,n}^i=(\widehat{\xi}_{p,n}^i(t)\,,\,\widehat{T}_{p-1,n}^i\leq t\leq\widehat{T}_{p,n}^i)\in E$$

The selection transition consists in randomly selecting a path-sequence

$$\zeta_n^i=(\xi_{0,n}^i,\cdots,\xi_{n,n}^i)$$

proportionally to its fitness

$$h_n(\xi_{0,n}^i,\cdots,\xi_{n,n}^i)=g_n(\xi_{n,n}^i)$$

The mutation stage consists in extending the selected paths according to an elementary $\mathcal{K}_{n+1}$-transition, that is

$$\zeta_{n+1}^i=((\xi_{0,n+1}^i,\cdots,\xi_{n,n+1}^i),\xi_{n+1,n+1}^i)=((\widehat{\xi}_{0,n}^i,\cdots,\widehat{\xi}_{n,n}^i),\xi_{n+1,n+1}^i)\in E_{n+1}=E_n\times E$$

where $\xi_{n+1,n+1}^i$ is a random variable with law $\mathcal{K}_{n+1}(\widehat{\xi}_{n,n}^i,\cdot)$. By a simple argument we see that the evolution associated with the end points of the paths

$$\xi_n=(\xi_{n,n}^1,\cdots,\xi_{n,n}^N)\quad\text{and}\quad\widehat{\xi}_n=(\widehat{\xi}_{n,n}^1,\cdots,\widehat{\xi}_{n,n}^N)\in E$$

coincide with the genetic algorithms described in Section 3. We conclude that the former path-particle Markov chain models the evolution in time of the corresponding genealogical

trees. For each time $n < \tau^N$ we denote by μ_n^N and $\widehat{\mu}_n^N$ the particle density profiles associated with the ancestor lines of this genealogical tree based algorithm

$$\mu_n^N = \frac{1}{N}\sum_{i=1}^N \delta_{(\xi_{0,n}^i,\cdots,\xi_{n,n}^i)} \quad \text{and} \quad \widehat{\mu}_n^N = \frac{1}{|I_n^N|}\sum_{i\in I_n^N} \delta_{(\xi_{0,n}^i,\cdots,\xi_{n,n}^i)}$$

with

$$I_n^N = \{1 \leq i \leq N \,:\, \xi_{n,n}^i(T_{n,n}^i) \in B_n\}$$

The asymptotic behavior of genealogical tree based algorithm has been studied in [8] in the context of strictly positive potentials and further developped in [4] for non negative ones. In our context the path-version of the $\mathbb{L}_p$-mean error estimates presented in theorem 2 can be stated as follows.

Theorem 5 *For any $p \geq 1$, $0 \leq n \leq m+1$ and any test function $f_n \in \mathcal{B}_b(E_n)$, with $\|f\| \leq 1$ we have*

$$(\mathbb{E}|\,\widehat{\mu}_n^N(f_n)\,1_{(\tau^N > n)} - \mathbb{E}(f_n([X_t\,,\;\; 0\leq t\leq T_n])\mid T_n \leq T)\,|^p)^{1/p} \leq a_p\, b_n/\sqrt{N}$$

for some finite constant $a_p < \infty$ which only depend on the parameter p and some finite constant $b_n < \infty$ which depends on the time parameter n.

Following the observations given the end of the previous section let us choose a collection of times $u_1 > 0$,..., $u_n > 0$. Let $f_n^{(u)}$, $u = (u_1, \cdots, u_n)$, be the test function on E_n defined by

$$f_n^{(u)}(x_0,\cdots,x_n) = f^{(u_1)}(x_1)\cdots f^{(u_n)}(x_n)$$

with $f^{(u_p)}$ defined in (3.6). In this situation we have

$$\mu_n(f_n^{(u)}) = \mathbb{P}(T_1 - T_0 \leq u_1, \cdots, T_n - T_{n-1} \leq u_n \mid T_n \leq T)$$

The particle approximations consists in counting at each level $1 \leq p \leq n$ the proportion of ancestral lines having succeeded to pass the p-th levels in time u_p.

In Figure 1 we illustrate the genealogical particle model associated with a particle X evolving in a pocket $C \subset S$ containing four "hard obstacles" R. We associate to a given stratification of the pocket C

$$R \subset C_0 \subset C_1 \subset C_2$$

the sequence of exit levels

$$B_0 = S \setminus R \supset B_1 = S \setminus C_0 \supset B_2 = S \setminus C_1 \supset B_3 = S \setminus C_2$$

The desired target set here is $B = B_3$.

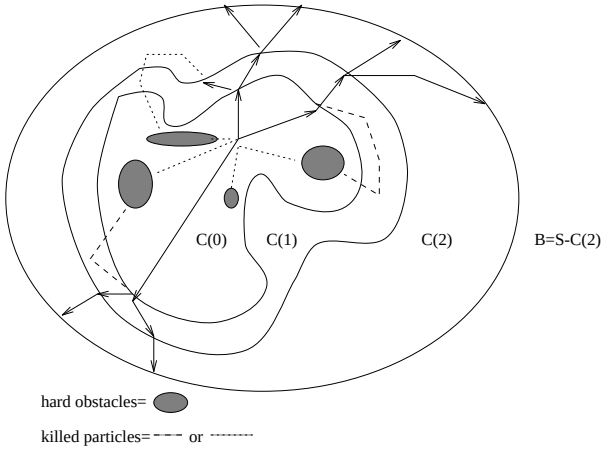

Figure 4.arabic@figure: Genealogical model, [exit of C(2) before killing] (N=7)

In Figure 2 we illustrate the genealogical particle model for a particle X evolving in a set $A \subset S$ with recurrent subset $R = S \setminus A$. To reach the desired target set B_4 the process need to pass the sequence of levels

$$B_0 \supset B_1 \supset B_2 \supset B_3 \supset B_4$$

5 Discussion

In this section we will discuss some practical aspects of the proposed method and compare it with the main other algorithms in the literature for the same purpose, that is *importance sampling (IS)* and *splitting*.

First of all, when IS already gives very good results, then very likeky that it is not necessary to find something else. One good feature of IS is to give i.i.d. sequences, which are quite simple to analyze. Very often the proposition distribution is chosen using large deviation arguments, at least in the case of static problems, see for instance [3]. But clearly it is not always obvious how to design an IS procedure for a given problem, especially for dynanical models such as Markov processes. Though in some very important practical problems, it may be quite easy to find a sequence of nested sets containing the rare event. In such cases, it is then appealing to use some splitting technique.

So let us focus now on splitting. Our main point here is that our algorithm has the same application domain as splitting, but performs better with virtually no additional cost. Let

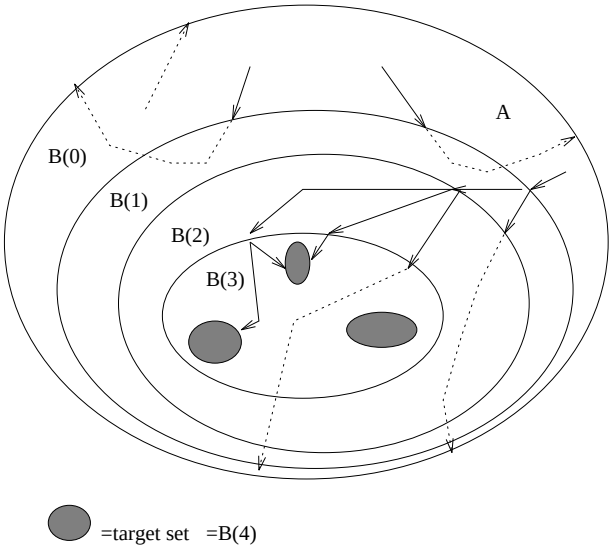

Figure 4.arabic@figure: Genealogical model, [ballistic regime, target B(4)] (N=4)

us consider a simplified framework: assume the Markov process is in one dimension, and that we have managed to set the levels such that all the probabilities $\mathbb{P}(T_{q+1} \leq T | T_q \leq T)$ are equal to the same P. For the splitting algorithm, assume all the branching rates are $1/P$. This is an optimal setting for the splitting algorithm as shown in [10]. In this case, the variance of the splitting astimator is:

$$n \frac{1-P}{P} \mathbb{P}(T_n \leq T)^2,$$

which is the same as the asymptotic variance of our algorithm as given by theorem 3. This means that the particle method performs (asymptotically) just as well as the splitting with optimal branching rates. So we have a method close to splitting, but with less parameters to tune, and still with the same accuracy. Moreover, the complexity is the same, the only added work is to randomly choose which particles have offsprings in the selection step, which is negligeable compared to the simulation of the trajectories. Note that both are much better than naive Monte-Carlo which have in this case a variance equal to:

$$\mathbb{P}(T_n \leq T)(1 - \mathbb{P}(T_n \leq T)).$$

It is also worth noting that the n factor in the asymptotic variance does not mean that the variance increases with n. For a given problen, the rare event probability is fixed, so the

level crossing probability is close to $P \simeq \mathbb{P}(T_n \leq T)^{\frac{1}{n}}$, and we have

$$\begin{aligned} n\frac{1-P}{P} &\simeq n(\exp[-\frac{1}{n}\log \mathbb{P}(T_n \leq T)] - 1) \\ &\simeq -\log \mathbb{P}(T_n \leq T) + \frac{1}{2n}\log^2 \mathbb{P}(T_n \leq T) + o(\frac{1}{n^2}), \end{aligned}$$

which means that as n goes to ∞, the variance is decreasing to $-\log[\mathbb{P}(T_n \leq T)]\ \mathbb{P}(T_n \leq T)^2$.

In practical applications the best is sometimes a pragmatic approach combining IS and our algorithm. Let us mention in this case [12], with numerical simulations of a hybrid model, which can be considered as a toy model for those used in Air Traffic Management. The theoretical study of this promissing approach is still to be done.

6 Numerical example : application to the Ornstein-Uhlenbeck process

We will show in this section how the previous method to simulate rare events works in a simple case. Although this is clearly a toy model, it allows us to check the method accuracy on the computation of some quantities that have formal rigorous expressions. Moreover, this process having simple Gaussian increments, there is no numerical error due to discretization scheme.

The process X is taken to be the 1-D Ornstein-Uhlenbeck process, i.e. the solution of the SDE

$$dX_t = -a\, X_t\, dt + \sigma\, \sqrt{2a}\, dW_t\ ,$$

where a and σ are strictly positive constants and W the standard Brownian motion in $\mathbb{R}$. The recurrent set R is chosen as $(-\infty, b^-]$, and then the process X is started at some $x_0 \in A = (b^-, +\infty)$. Given some $b^+ > x_0$, we set the target $B = [b^+, +\infty)$. It is clear that if we take b^+ large enough, the probability to hit the target can be made arbitrarily small. Let us denote by τ the stopping time

$$\tau = \inf\{t > 0\ :\ X_t \notin (b^-, b^+)\}\ .$$

In order to check the method, we will compute $\mathbb{E}[\tau \mid X_\tau = b^+]$ using both a Monte–Carlo method based on our rare event analysis approach and the theoretical expression. From [2] we have

$$\mathbf{L}(\alpha) = \mathbb{E}_{x_0}[e^{-\alpha\tau}\mathbf{1}(X_\tau = b^+)] = \frac{S(\frac{\alpha}{a}, \frac{x_0}{\sigma}, \frac{b^-}{\sigma})}{S(\frac{\alpha}{a}, \frac{b^+}{\sigma}, \frac{b^-}{\sigma})}$$

where S is a special function to be defined in the sequel. Using the derivative of the Laplace transform we get

$$\mathbb{E}[\tau \mid X_\tau = b^+] = -\frac{1}{\mathbb{P}(X_\tau = b^+)} \left. \frac{d\mathbf{L}(\alpha)}{d\alpha} \right|_{\alpha=0} . \tag{6.1}$$

The probability in the denominator is given by

$$\mathbb{P}(X_\tau = b^+) = \frac{u(x_0) - u(b^-)}{u(b^+) - u(b^-)}, \tag{6.2}$$

where the function u (the scale function of the process) is in our case a primitive of $u'(x) = \exp\{\frac{x^2}{2\sigma^2}\}$. This function u is then easily computed using any standard numerical integration routine. The derivative of $\mathbf{L}$ is more tricky. First we write the expression of the function S, for any real x and y, and $\nu > 0$,

$$S(\nu, x, y) = \frac{\Gamma(\nu)}{\pi} \, e^{\frac{1}{4}(x^2+y^2)} \, [\, D_{-\nu}(-x) D_{-\nu}(y) - D_{-\nu}(x) D_{-\nu}(-y) \,] \, ,$$

where the functions D are the parabolic cylinder functions defined by

$$\begin{aligned} D_{-\nu}(x) \quad = \quad & e^{-\frac{1}{4}x^2} \, 2^{-\frac{\nu}{2}} \, \sqrt{\pi} \, \Big[\, \frac{1}{\Gamma(\frac{1}{2}\,(\nu+1))} [\, 1 + \sum_{k=1}^{\infty} \frac{\nu(\nu+2)\dots(\nu+2k-2)}{3 \cdot 5 \cdots (2k-1)\; k!} (\tfrac{1}{2}\, x^2)^k \,] \\ & - \frac{x\sqrt{2}}{\Gamma(\frac{1}{2}\,\nu)} [\, 1 + \sum_{k=1}^{\infty} \frac{(\nu+1)(\nu+3)\cdots(\nu+2k-1)}{3 \cdot 5 \cdots (2k+1)\; k!} (\tfrac{1}{2}\, x^2)^k \,] \, \Big] \; . \end{aligned}$$

These functions are computed using the numerical method and the source code provided in [11]. Now we still need to compute the derivative in equation (6.1). We did not want to derive formaly this quite complicated expression, and used instead a numerical appriximation from a local rate of variation:

$$\left. \frac{d\mathbf{L}(\alpha)}{d\alpha} \right|_{\alpha=0} \simeq \frac{\mathbf{L}(2\varepsilon) - \mathbf{L}(\varepsilon)}{\varepsilon},$$

where $\varepsilon > 0$ is chosen small enough.

Now we explain how the Monte-Carlo computation was carried out. For the decreasing sequence of Borel sets $\{B_j \, , \, j = 1 \dots M\}$ we chose an increasing squence of real numbers $\{b_j \, , \, j = 1 \dots M\}$, with $b^- < b_1 < \cdots < b_M < b^+$ and take $B_j = (b_j, +\infty)$. In our special case, we can choose the probability for a particle started at b_j to reach b_{j+1}, and compute both he number of levels and each level accordingly. If we take these probability equal for all j to say p, then

$$M = \lfloor \frac{\log \mathbb{E}[\tau \mid X_\tau = b^+]}{\log p} \rfloor.$$

Alternatively we can choose M and compute p. Note that the probability of the N–particle cloud to be killed before reaching b^+ is $1-(1-(1-p)^N)^M$ which can be small even with a small number N of particles when p is say larger than $1/2$. From this we see that a good strategy is to make many runs of our algorithm on a small number of particles, instead of only a few runs on a large number of particles (on the same run, all the generated trajectories are obviously strongly correlated). All the corresponding values b_j are easily computed using expressions as the one in equation (6.2).

In Figure 6.arabic@figure we see the expectation $\mathbb{E}[\tau \mid X_\tau = b^+]$ as a function of b^+, with $b^- = 0$. The blue curve is the numerically computed theoretical value, and the red curve is the Monte–Carlo simulation result, with 880 runs of 8 particles each. The parameters of the Ornstein-Uhlenbeck process are $a = 0.1$, $\sigma\sqrt{2a} = 0.3$ and $x_0 = 0.1$. The largest value of b^+ was 4.0. This means that the probability for the process started at $x_0 = 0.1$ to reach the desired level is approximatly 1.6460×10^{-08}, so there is no way of simulating trajectories by the naive approach.

Another examples of rare events for diffusions may be found in Aldous [1], which presents a Poisson clumping heuristic as well as numerous examples and references. For instance, following [1, Section I11], let consider a diffusion in $\mathbb{R}^d$ starting from 0 with drift $\mu(x) = -\nabla H(x)$ and variance $\sigma(x) = \sigma_0 I$. Suppose H is a smooth convex function attaining its minimum at 0 with $H(0) = 0$ and such that $H(x) \to \infty$ as $|x| \to \infty$. Let B be a ball with center at 0 with radius r, where r is sufficiently large that $\pi(B^c)$ is small, where π is the stationary distribution

$$\pi(x) = c \exp\{ \frac{-2\,H(x)}{\sigma_0^2} \} \approx (\sigma_0^2 \pi)^{-d/2} |Q|^{1/2} \exp\{ \frac{-2\,H(x)}{\sigma_0^2} \} \ ,$$

where

$$Q = \left(\frac{\partial^2 H}{\partial x_i \partial x_j}(0) \right)_{i,j \geq 1} \ .$$

We want an estimation of the first exit time from the ball B. There are two qualitatively situations : radially symmetric potentials ($H(x) = h(|x|)$) and non-symmetric potentials. We presents here only the second one, by assuming that H attains its minimum, over the spherical surface ∂B, at a unique point $z_0 = (r, 0, 0, \cdots)$. Since the stationary distribution decreases exponentially fast as H increases, we can suppose that exits from B will likely occur near z_0 and then approximate T_B by T_F, the first hitting time on the $(d-1)$-dimensional hyperplane F tangent to B at z_0. The heuristic used in [1] gives that T_B is approximatively exponentially distributed with mean $(\pi_F\,|\nabla H(z_0)|)^{-1}$, where π_F designed the restriction of

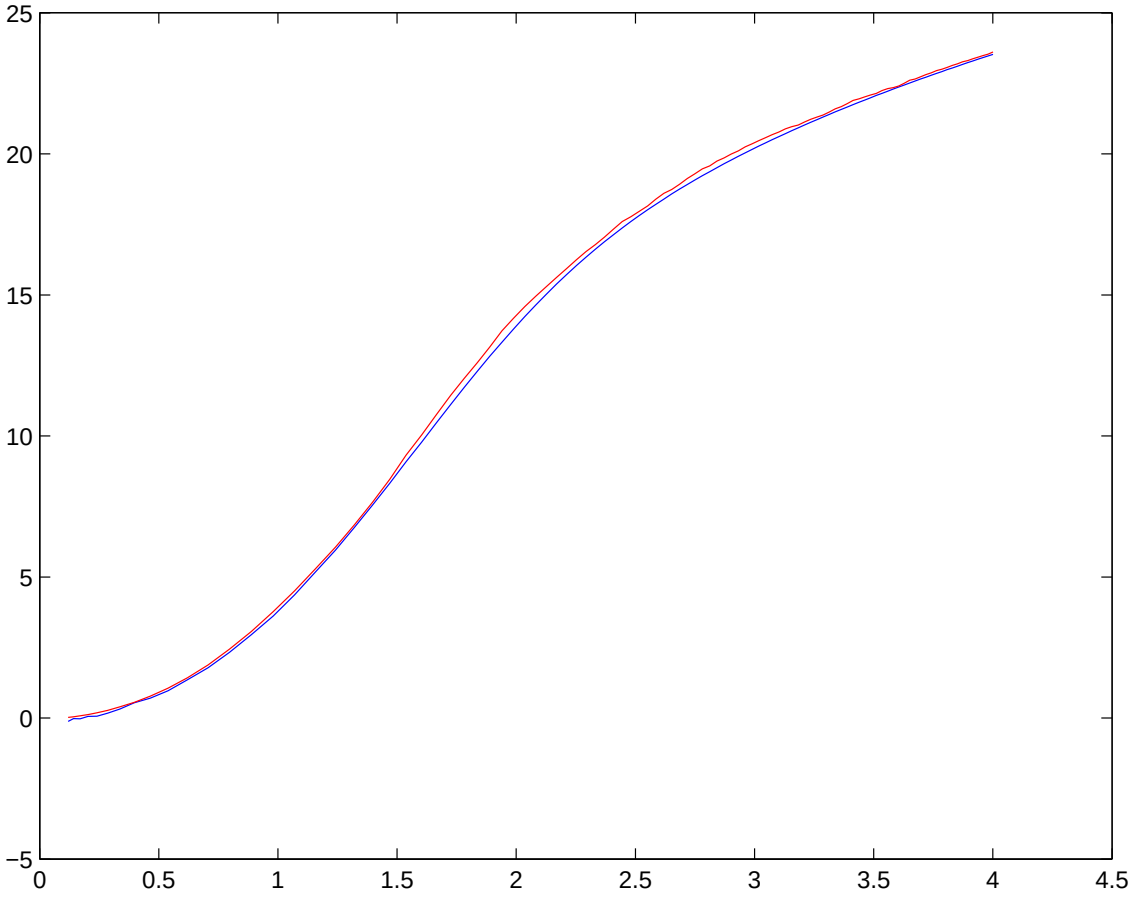

Figure 6.arabic@figure: Theoretical and Monte–Carlo mean conditional stopping times

the measure π to F. We obtain

$$\begin{aligned}\pi_F &\approx (\sigma_0^2\pi)^{-d/2}\,|Q|^{1/2}\,\exp\{-\frac{2\,H(z_0)}{\sigma_0^2}\}\,\int_F \exp\{-\frac{2\,(H(x)-H(z_0)}{\sigma_0^2}\}\,dx\\ &\approx (\sigma_0^2\pi)^{-1/2}\,|Q|^{1/2}\,|Q_1|^{-1/2}\,\exp\{-\frac{2\,H(z_0)}{\sigma_0^2}\}\,,\end{aligned}$$

where

$$Q_1 = \left(\frac{\partial^2 H}{\partial x_i \partial x_j}(z_0)\right)_{i,j\geq 2}\,.$$

Thus

$$\mathbb{E}(T_B) \approx \sigma_0 \frac{\pi^{1/2}|Q|^{-1/2}|Q_1|^{1/2}}{-\dfrac{\partial H}{\partial x_1}(z_0)}\;\exp\{\frac{2\,H(z_0)}{\sigma_0^2}\}\,.$$

The simplest concrete example is the Ornstein-Uhlenbeck process in wich $H(x) = \frac{1}{2}\sum \rho_i x_i^2$ whith $0 < \rho_1 < \rho_2 < \cdots$. Here H has two minima on ∂B, at $\pm z_0 = \pm(r, 0, 0, \cdots)$ and so the mean exit time is

$$\mathbb{E}(T_B) \approx \frac{1}{2}\sigma_0\pi^{1/2}(\prod_{i\geq 2}\rho_i/\prod_{i\geq 1}\rho_i\,)\;\rho_1^{-1}r^{-1}\exp\{\frac{\rho_1 r^2}{\sigma_0^2}\}\,.$$

To adapt this example to the formalism, introduced previously, we slightly modify it by considering the first exit time from the ball B before reaching a little ball B_ε centered at 0 with radius ε small. Thus, we suppose that $\mathbb{R}^d$ is decomposed into two separate regions B^c and B and that the process X evolves in B starting from outside B_ε, but near from ∂B_ε. The process will be killed as soon as it hits ∂B_ε. By considering a particle system algorithm and a genealogical model, an estimation of the first exit time before returning in the neighbourhood of the origin and of the distribution of the process during its excursions should be obtained.

It will also be interesting to study the Kramers equation

$$\begin{cases} dX_t &= V_t\,dt\\ dV_t &= -H'(X_t)\,dt - \gamma\,V_t\,dt + \sqrt{2\,\gamma}\,dB_t\end{cases}$$

In [1, Section I13], the heuristic may be applied for small and large coefficients γ, but it is a hard problem to say which of these behaviors dominates in a specific non-asymptotic case, hence the simulation approach.

References

[1] D. Aldous. *Probability Approximations via the Poisson Clumping Heuristic*, volume 77 of *Applied Mathematical Sciences*. Springer Verlag, 1989

[2] A.N. Borodin and P. Salminen. *Handbook of Brownian Motion — Facts and Formulae*. Probability and its Applications. Birkhäuser, Basel, 1996.

[3] J.A. Bucklew, *Introduction to rare event simulation*, Springer Series in Statistics, Springer-Verlag, New York, 2004.

[4] F. Cérou, P. Del Moral and F. LeGland, *On genealogical trees and Feynman–Kac models in path space and random media*, Preprint 2002.

[5] P. Del Moral, J. Jacod and Ph. Protter, The Monte-Carlo method for filtering with discrete-time observations. *Probability Theory and Related Fields*, vol. 120, no. 3, 346–368 (2001).

[6] P. Del Moral, M.A. Kouritzin and L. Miclo, On a class of discrete generation interacting particle systems. *Electronic Journal of Probability*, vol. 6, no. 16, 1–26, 2001.

[7] P. Del Moral and L. Miclo, *Branching and interacting particle systems approximations of Feynman–Kac formulae with applications to non–linear filtering,* Séminaire de Probabilités XXXIV. Lecture Notes in Mathematics No. 1729, J. Azéma, M. Émery, M. Ledoux, M. Yor (Eds.), pp. 1–145, 2000.

[8] P. Del Moral and L. Miclo, Genealogies and increasing propagations of chaos for Feynman–Kac and genetic models. *Annals of Applied Probability*, vol. 11, no. 4, 1166–1198 (2001).

[9] P. Glasserman, P. Heidelberger, P. Shahabuddin and T. Zajic, Multilevel splitting for estimating rare event probabilities. *Operations Research*, vol. 47, no. 4, 585-600, July-August 1999.

[10] A. Lagnoux. Rare event simulation. *PEIS*, vol. 20, no. 1, 45–66 (2006).

[11] Shan-jie Zhang and Jianming Jin, *Computation of Special Functions*. John Wiley & Sons, New York, 1996.

[12] J. Krystul and H.A.P. Blom, *Sequential Monte Carlo Simulation of Rare Event Probability in Stochastic Hybrid Systems*, in Proceedings of the 16th IFAC World Congress, 2005.

[13] B. Tuffin and K.S. Trivedi, *Implementation of importance splitting techniques in stochastic Petri net package* TOOLS'00 : Computer Performance Evaluation. Modelling Techniques and Tools. Lecture in Computer Science No. 1786, B. R. Harverkort, H. C. Bohnenkamp, C. U. Smith, pp. 216–229, 2000.

[14] M. Villén-Altamirano and J. Villén-Altamirano, *RESTART: A method for accelerating rare event simulations.* In 13th Int. Teletraffic Congress, ITC 13 (Queueing, Performance and Control in ATM), pp. 71–76, Copenhagen, Denmark, 1991

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399